

V Section States (Section)





APPROXIMATION FOR BAYESIAN ABILITY ESTIMATION

Robert K. Tsutakawa and Michael J. Soltys

Mathematical Sciences Technical Report No. 134 February 1987

Department of Statistics University of Missouri Columbia, MO 65211





Prepared under contract No. N00014-85-K-0113, NR 150-535 with the Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research

Approved for public release: distribution unlimited.
Reproduction in whole or in part is permitted for any purpose of the United States Government

	STASSIFICATION O		PAGE	AN-AI	79/03				
				DOCUMENTATIO				rm Approved AB No. 0704-0188	
1a REPORT SECURITY CLASSIFICATION Unclassified					16 RESTRICTIVE MARKINGS				
2a SECURITY	CLASSIFICATIO	N AUT	HORITY		3 DISTRIBUTION/AVAILABILITY OF REPORT Approval for public release; distribution				
26 DECLASSIFICATION / DOWNGRADING SCHEDULE					unlimited. Reproduction in whole or part i permitted for any purpose of the US governme				
4 PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report No. 134					5 MONITORING	5 MONITORING ORGANIZATION REPORT NUMBER(S)			
6a NAME OF PERFORMING ORGANIZATION Department of Statistics				6b OFFICE SYMBOL (If applicable)	Personnel	7a NAME OF MONITORING ORGANIZATION Personnel and Training Research Programs Office of Naval Research			
University of Missouri 6c. ADDRESS (City, State, and ZIP Code)						7b. ADDRESS (City, State, and ZIP Code)			
Columbia, MO 65211					Arlington, VA 22217				
8a. NAME OF FUNDING/SPONSORING ORGANIZATION				8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-85-K-0113				
8c. ADDRESS	(City, State, and	J ZIP Co	de)	4 		10 SOURCE OF FUNDING NUMBERS			
					PROGRAM ELEMENT NO 61153N	PROJECT NO: RR042-04	TASK NO: 042-04-1	WORK UNIT ACCESSION NO. 4421-535	
App 12 PERSONAL Tsu	L AUTHOR(S) takawa, R	n for	Bayesia	an Ability Estin	J.				
13a TYPE OF REPORT 13b TIME C Technical FROMB5J				OVERED 14 DATE OF REPORT (Year, Montant and			h, Day) 15 PAGE COUNT 32		
16 SUPPLEME	ENTARY NOTAT	TION							
17 COSATI CODES				18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)					
FIELD	GROUP	SUE	8-GROUP	Ability estimation, Bayesian approximation, Item response theory					
An apprility particle post on the pon the i	oximation ameter in erior mean osterior n tem parame	is p an i n and mean eters	proposed item resp I covaria of a Tay	and identify by block for for the poster: conse model. The ance matrix of a ylor series appr	ior mean and he procedure item parameto roximation to	assumes the	at approx: ilable. : rior mean	imations to It is based	
	TION / AVAILABI SIFIED/UNLIMIT			RPT DTIC USERS		CURITY CLASSIFICA	TION		

DD Form 1473, JUN 86

22a NAME OF RESPONSIBLE INDIVIDUAL Dr. Charles Davis

Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

22b TELEPHONE (Include Area Code) 22c OFFICE SYMBOL (202) 696–4046

APPROXIMATION FOR BAYESIAN ABILITY ESTIMATION

ROBERT K. TSUTAKAWA

and

MICHAEL J. SOLTYS

University of Missouri

QUALITY (IIS) - GTE 1

Correspondence to be handled by:

Robert K. Tsutakawa Department of Statistics University of Missouri 316 Math Sciences Columbia, MO 65211 314-882-6879

APPROXIMATION FOR BAYESIAN ABILITY ESTIMATION

KEY WORDS: Ability estimation, Bayesian approximation, binary variables, item response, two-parameter logistic.

Abstract

An approximation is proposed for the posterior mean and standard deviation of the ability parameter in an item response model. The procedure assumes that approximations to the posterior mean and covariance matrix of item parameters are available. It is based on the posterior mean of a Taylor series approximation to the posterior mean conditional on the item parameters. The method is illustrated for the two-parameter logistic model using data from an ACT math test with 39 items. A numerical comparison with the empirical Bayes method shows that the point estimates are very similar but the standard deviations under empirical Bayes are about two percent smaller than those under Bayes. The effect of sample size is demonstrated by illustrating the increase in the standard deviations for a smaller data set.

Ability Estimation

We consider estimating the ability of an individual based on the person's dichotomous responses to a set of test items whose characteristics are partially known through responses to the same set from other individuals belonging to the same population. Each item is characterized by an item response function which defines the probability of a correct response to the item by an individual with ability θ . We assume such a function to be known except for some parameter ξ and use the notation $P(y|\theta,\xi)$ to represent the probability of correct response (y=1) or incorrect response (y=0) by an individual with real valued ability θ to an item with parameter ξ . An example of such a function is the two-parameter logistic model defined by

$$P(y|\theta,\alpha,\beta) = \frac{\exp\{y\alpha(\theta-\beta)\}}{1+\exp\{\alpha(\theta-\beta)\}},$$
 (1)

y = 0,1, where $-\infty$ < θ < ∞ , ξ = (α,β) is two dimensional with α > 0 and $-\infty$ < β < ∞ .

We assume the calibration of the test is based on responses to k items by a group of n individuals representative of some target population for which abilities are to be measured. We let $y_{ij} = 0$ or 1 according as the response by examinee i to item j is incorrect or correct. We assume conditional independence among the responses so that the joint probability of the nxk matrix y of responses y_{ij} is given by

$$P(y|\theta,\xi) = \prod_{ij} (y_{ij}|\theta_{i},\xi_{j})$$
 (2)

where $\theta = (\theta_1, \dots, \theta_n)$ represents the abilities of the n individuals and $\xi = (\xi_1, \dots, \xi_k)$ the parameters of the k items.

One standard method used in calibration is to obtain the joint maximum likelihood estimate $(\hat{\mathbb{A}}_L,\hat{\mathbb{A}}_L)$ of (θ,ξ) using a procedure such as LOGIST (Wingersky, Barton, and Lord, 1982). The calibrated items are then used to measure the ability θ for a new individual with item responses $\mathbf{x}=(\mathbf{x}_1,\ldots,\mathbf{x}_k)$ by finding the value of θ which maximizes the probability

$$P(\mathbf{x} \mid \theta, \hat{\xi}_{Lj}) = \Pi P(\mathbf{x}_{j} \mid \theta, \hat{\xi}_{Lj}).$$

This procedure is straight forward and relatively easy to implement, however it is known to be biased outwards for extreme values of θ (Lord, 1983a), is subject to an occasional nonexistence of a solution, and fails to provide a good measure of uncertainity in the estimated θ .

A related empirical Bayes procedure is to first assume a prior distribution π of θ and estimate ξ by marginal maximum likelihood. Then derive the posterior mean of θ conditionally on assumed to equal the marginal maximum likelihood estimate. The resulting estimate of θ is an empirical Bayes (Cox and Hinkley, 1974) estimate and has been demonstrated on different models by Bock and Aitkin (1981) and Rigdon and Tsutakawa (1983), among others.

The Bayesian approach, which we adopt here, assumes prior distributions on both θ and ξ and uses an approximate posterior mean and variance of θ to make inferences regarding the unknown θ . The approximation for the posterior mean $E(\theta \mid X)$ is based the posterior expectation of a Taylor series approximation of the

Ability Estimation

conditional posterior expectation $E(\theta \mid \mathbf{x}, \boldsymbol{\xi})$. When the posterior distribution of $\boldsymbol{\xi}$ is normal, it reduces to Lindley's (1980) approximation whose general form requires the computation of a large number of 3rd partial derivatives of the loglikelihood function. We will first discuss our method when the individual whose θ is being measured, is a member of the calibrating set, e.g. $\theta = \theta_1$. We then show how the same approximation can be modified and used on a new individual from the same target population.

Approximate Posterior Mean and Variance of θ

Assume that θ_1,\dots,θ_n are independent and identically distributed (iid) according to some known prior $p(\theta)$. Also assume that ξ_1,\dots,ξ_k are independently distributed according to some known priors $p(\xi_1),\dots,p(\xi_k)$ and independent of θ_1,\dots,θ_n . Then the marginal posterior pdfs of ξ and θ are given by

$$p(\xi|\hat{y}) = p(\xi) = \int_{\hat{i}} \frac{\pi}{i} P(y_{ij}|\theta_i,\xi_j)p(\theta_i)d\theta_i$$
(4)

and

$$p(\theta | \hat{y}) \propto p(\theta) \prod_{j} \prod_{i} P(y_{ij} | \theta_{i}, \xi_{j}) p(\xi_{j}) d\xi_{j}.$$
 (5)

As shown in Tsutakawa and Lin (1986) the EM algorithm can be used to compute the posterior mode $\hat{\xi}$ of ξ and the posterior covariance of ξ can be approximated by Σ , the inverse of the negative Hessian of the log posterior evaluated at $\hat{\xi}$. Due to the symmetry between (4) and (5), it may appear that similar methods may be adopted to compute the posterior mode of θ . However, the analogous approach presents serious numerical problems since the integrals in (5) are multiple integrals in contrast to the single integrals in (4).

Instead we shall try to derive the posterior expectation of θ by approximating the integral in the expression,

$$E(\hat{\theta} | \hat{\mathbf{y}}) = \int E(\hat{\theta} | \hat{\mathbf{y}}, \hat{\xi}) p(\hat{\xi} | \hat{\mathbf{y}}) d\hat{\xi}, \qquad (6)$$

where $\mathrm{E}(\underbrace{\theta} \mid \underline{y}, \underline{\xi})$ is the posterior expectation of $\underbrace{\theta}$ conditionally on $\underline{\xi}$, which is used by the empirical Bayes approach upon substituting $\underline{\xi}$ with its marginal maximum likelihood estimate.

Let $w(\xi) = E(\theta_i | y_i, \xi)$ and let w, w_r , w_{rs} and w_{rst} denote the values of $w(\xi)$ and its first three derivatives all evaluated at ξ . Then according to Lindley (1980), under regularity conditions (6) may be approximated to order $O(n^{-1})$ by

$$w + \frac{1}{2} \sum_{i} w_{i} rs^{T} rs^{+\frac{1}{2}} \sum_{i} \Lambda_{rst} w_{i}^{T} rs^{T} tu$$
 (7)

where $\Lambda_{\rm rst}$ is the third partial of the log posterior evaluated at $\xi=\dot{\xi}$ and $\tau_{\rm rs}$ are the elements of ξ . The last term accounts for the skewness of the posterior distribution of ξ and if, in particular, the posterior is normal, $\Lambda_{\rm rst}$ vanish(Anderson, 1958, p. 39) and (7) reduces to

$$m = w + \frac{1}{2} \sum_{\mathbf{w}_{\mathbf{r}\mathbf{S}}} \tau_{\mathbf{r}\mathbf{S}}. \tag{8}$$

A heuristic justification for (8) can be given by considering the posterior expectation of the third order Taylor series approximation to $E(\theta_1 | y_1, \xi)$ about $\xi = \frac{\hat{\xi}}{\xi}$. To simplify the notation represent the components of ξ and ξ by (u_1, \dots, u_q) and (u_1, \dots, u_q) . Then the approximation is given by

$$w(\xi) = w + \sum_{r} w_{r} (\mu_{r} - \mu_{r})$$

$$+ \frac{1}{2\Sigma\Sigma} w_{rs} (\mu_{r} - \mu_{r}) (\mu_{s} - \mu_{s})$$

$$+ \frac{1}{6} \sum_{rst} w_{rst} (\mu_{r} - \mu_{r}) (\mu_{s} - \mu_{s}) (\mu_{t} - \mu_{t}). (9)$$

If $p(\xi|y)$ is normal with mean ξ and covariance matrix (σ_{rs}) , then the third moments of ξ vanish and

$$\int \mathbf{w}(\xi) \, \mathbf{p}(\xi \, | \, \mathbf{y}) \, \mathrm{d}\xi$$

$$\stackrel{\cdot}{=} \mathbf{w} + \frac{1}{2} \, \Sigma \mathbf{w}_{rs} \, \sigma_{rs}. \qquad (10)$$

Upon replacing ((σ_{rs})) by its estimate ((τ_{rs})) we have (8). Mosteller and Wallace (1964, p. 151) refer to (10) as the standard approximation for the case of known means and covariances and point out the need to consider the bias $\frac{\Lambda}{\xi} - E(\xi | \underline{y})$ and third derivatives when the distribution is nonsymmetric.

The posterior variances and covariances of $\frac{1}{2}$ may be similarly approximated by first obtaining an approximation to $E(\theta_i \theta_j | \underline{y})$ by repeating (8) with $w(\xi) = E(\theta_i \theta_j | \underline{y}_i, \underline{y}_j, \underline{\xi})$ and then making the appropriate substitutions in $E(\theta_i \theta_j | \underline{y}) - E(\theta_i | \underline{y}) E(\theta_j | \underline{y})$. The approximate variances, s^2 , and covariances can be helpful in assessing the uncertainty of a particular individual's ability or in comparing the abilities of two individuals.

Suppose we now have the n+1st or a new individual, with unknown θ , whose response to the same k items is $\underline{x} = (x_1, \dots, x_k)$. We can repeat the entire procedure including the updating of (ξ, ξ) after replacing \underline{y} with $(\underline{y}, \underline{x})$. However this would be quite costly for routine evaluations. It would be much simpler retaining (ξ, ξ) from the first n individuals and using \underline{x} in $w(\xi) = E(\theta | \underline{x}, \xi)$ only. When n is moderately large, the change in (m, s^2) through the updated (ξ, ξ) by the addition of a single individual would be quite negligible relative to the variability in $E(\theta | \underline{x}, \xi)$ or $E(\theta^2 | \underline{x}, \xi)$ due to different values of \underline{x} .

Computational Details for the Two-parameter Logistic Model

In this section we present computational expressions for the evaluation of w and w_{rs} in the case of the two-parameter logistic model (1) with a N(0,1) prior on θ . A derivation of (ξ, Σ) based on the EM algorithm can be found in Tsutakawa and Lin (1986).

For this model we have

$$w(\xi) = E(\theta_{i} | Y_{i}, \xi)$$

$$= \int \theta_{i} p(\theta_{i} | Y_{i}, \xi) d\theta_{i}, \qquad (11)$$

where

$$p(\theta_{i}|y_{i},\xi) = \frac{p(\theta_{i}) \prod_{j=1}^{k} p(y_{ij}|\theta_{i},\alpha_{j},\beta_{j})}{\prod_{j=1}^{k} p(y_{ij}|\theta_{i},\alpha_{j},\beta_{j})}$$

and

$$p(\theta) = \frac{1}{\sqrt{2\pi}} \exp(-\theta^2/2), -\infty < \theta < \infty.$$

Following the notation in Tsutakawa (1984), for each i = 1,...,n and u = 1,...,k let

$$g_{1}(i,u,\theta) = \frac{\partial P(y_{iu} | \theta, \alpha_{u}, \beta_{u})}{\partial \alpha_{u}} / P(y_{iu} | \alpha, \alpha_{u}, \beta_{u}), \qquad (12)$$

$$g_2(i,u,\theta) = \frac{\partial P(y_{iu}|\theta,\alpha_u,\beta_u)}{\partial \beta_u} / P(y_{iu}|\theta,\alpha_u,\beta_u),$$

$$h_{11}(i,u,\theta) = \frac{\partial^{2} P(y_{iu} | \theta, x_{u}, \beta_{u})}{\partial \alpha_{u}^{2}} / P(y_{iu} | \theta, x_{u}, \beta_{u}),$$

$$h_{12}(i,u,\theta) = \frac{\theta^{2}P(y_{iu}|A, x_{u}, \beta_{u})}{\theta^{\alpha}u^{\theta}\beta_{u}}/P(y_{iu}|\theta, \alpha_{u}, \beta_{u}),$$

$$h_{22}(i,u,\theta) = \frac{\vartheta^{2}P(y_{iu}|\theta,\alpha_{u},\beta_{u})}{\vartheta\beta_{u}^{2}}/P(y_{iu}|\theta,\alpha_{u},\beta_{u}),$$

$$\phi_{\theta \mathbf{u}} = \{1 + \exp[-\alpha_{\mathbf{u}}(\theta - \beta_{\mathbf{u}})]\}^{-1}, \text{ and } \psi_{\theta \mathbf{u}} = 1 - \phi_{\theta \mathbf{u}}.$$

By taking derivatives it is readily shown that

$$\begin{split} g_{1}(i,u,\theta) &= (y_{1u} + \phi_{\alpha u})(\theta - \beta_{u}), \\ g_{2}(i,u,\theta) &= -x_{u}(y_{1u} - \phi_{\theta u}), \\ h_{11}(i,u,\theta) &= (y_{1u} - \phi_{\theta u})(\phi_{\alpha u} - \phi_{\alpha u})(\theta - \beta_{u})^{2}, \\ h_{12}(i,u,\theta) &= -(y_{1u} - \phi_{\theta u})(1 + x_{u}(\phi_{\alpha u} - \phi_{\theta u})(\theta - \beta_{u})), \end{split}$$
(13)

and

$$h_{22}(i,u,\theta) = (y_{iu} - \phi_{iu}) \pi_u^2 (\gamma_{iu} - \phi_{iu}).$$

(We note that there is a minus sign missing in the expression for h_{12} in Tsutakawa (1984)). Now define the following posterior expectations of the derivatives and their products by

$$\bar{g}_{s}(i,u) = \int_{-\infty}^{\infty} g_{s}(i,u,\theta) p(\theta | y_{i}, \xi) d^{\alpha},$$

$$\frac{1}{h_{st}}(i,u) = \int_{-\infty}^{\infty} h_{st}(i,u,\theta) p(\theta|y_i,\xi) d\theta,$$
(14)

$$\bar{d}_{st}(i,u,v) = \int_{-\infty}^{\infty} g_s(i,u,\theta)g_t(i,v,\theta)p(\theta|y_i,\xi)d\theta,$$

s,t = 1,2,u,v = 1,...,k,u \neq v, where p($\theta \mid y_i, \xi$) is defined by (11).

Now denote the second derivatives of w(ξ) evaluated at ξ = ξ by

$$w_{11}(u, v) = \frac{\partial w(\xi)}{\partial \alpha_{u}} \partial \alpha_{v},$$

$$w_{12}(u, v) = \frac{\partial w(\xi)}{\partial \alpha_{u}} \partial \beta_{v},$$

$$w_{21}(u, v) = \frac{\partial w(\xi)}{\partial \alpha_{u}} \partial \alpha_{v},$$
(15)

and

$$w_{22}(u,v) = \frac{\partial w(\xi)}{\partial \theta_u} \hat{\partial}_{v}.$$

These derivatives may be computed using the following expressions.

$$w_{st}(u,u) = E\{\theta h_{st}(i,u,\theta)\} + \{2\bar{g}_{s}(i,u)\bar{g}_{t}(i,u) - \bar{h}_{st}(i,u)\}E(4) - \bar{g}_{s}(i,u)E\{\theta g_{t}(i,u,\theta)\} - \bar{g}_{t}(i,u)E\{\theta g_{s}(i,u,\theta)\}$$
(16)

and, for $u \neq v$,

$$w_{st}(u,v) = \{2\bar{g}_{s}(i,u)\bar{g}_{t}(i,v) - \bar{d}_{st}(i,u,v)\}E(4)$$
(17)

+
$$E^{(i)}g_{s}(i,u,v)g_{t}(i,v,v)$$

- $\bar{g}_{t}(i,v)E^{(i)}g_{s}(i,u,v) - \bar{g}_{s}(i,u)E^{(i)}g_{t}(i,v,v)$,

where E denotes posterior expectation with respect to the density $p\left(\frac{x}{1},\frac{y}{2},\frac{x}{2}\right)$ defined by (11) when $\xi=\xi$.

The numerical evaluation of $w_{\mbox{st}}(u,v)$ requires numerically integrating a number of integrals of the type

$$\int_{-1}^{\infty} b(x_{1}) p(x_{1}) \frac{dx_{1}}{dx_{1}} P(y_{1}, x_{1}, x_{2}, x_{3}) dx_{1}, \qquad (18)$$

where the function $h(\frac{\pi}{i})$ varies from one integral to the next. The use of scaling for improved accuracy in the Gauss-Hermite approximation is discussed in Tsutakawa (1984).

Example

Our method will now be illustrated on a sample data set from an ACT math test previously used by Tsutakawa and Lin (1986). The thirty of responses on k=39 items by n=400 respondents.

sing their method, $(\frac{c}{2}, \frac{c}{2})$ were initially recomputed. The posterior means and standard deviations (m,s) were then some that the thirst 100 respondents and are plotted against the empirical minimum and mini

where $\hat{\xi}$ is the marginal maximum likelihood estimate of ξ] in Figures 1 and 2. We note that the means under the two procedures are almost intical but the standard deviations could be quite different.

In particular, posterior standard deviations are 1.5 to 2.5 percent smaller under empirical Bayes than approximate Bayes.

ne explanation for this is that the empirical Bayes approach does not account for the uncertainity in $\frac{x}{z_L}$ (Deeley and Lindley 1981).

A plot of s vs.m in Figure 3 shows that the larger s is associated with the larger 'm', suggesting the difficulty of estimating extreme ability values based on the given test. The quadratic relationship seen in the plot suggests that, for routine use, s may be expressed as a function of m and need not be computed for all individuals.

Effect of Sample Size n

An interesting conjecture involves the effect of the sample size n on our Bayesian estimates. Although the posterior moments of ability are calculated separately for each individual they depend on the values of (ξ, Σ) which are estimated from a response matrix which may contain the particular individual's response. It is natural to believe that as the size of the response matrix decreases, the posterior standard deviation of an individual's ability would increase, as there is more uncertainty in the estimated (ξ, Σ) .

We previously found ability estimates for the first 100 examinees in a 400 x 39 response matrix using (ξ, Σ) which were derived from the responses of all 400 examinees. To investigate our conjecture we reestimated the posterior moments of ability for the same set of 100 examinees, using (ξ, Σ) derived from the first 100 examinees only.

The initial comparison produced an interesting result. From individual to individual it was found that the posterior standard deviation of ability was usually smaller than the corresponding standard deviation based on n=400. We do not feel however that our initial conjecture is untrue as the following discussion shows.

It is well known that the parameters of the two-parameter logistic model are not unique. The item parameters which were estimated from the two samples are quite different although the items are exactly the same. In particular, the estimates of $\alpha_{\mbox{\scriptsize j}}$

were larger for n = 100. This is consistent with Lord (1983b) who shows that the MLE of the discrimination parameter is positively biased, particularly for small n. (Although we are not using MLE's here, our prior is relatively flat.) Intuitively, larger values of the discrimination parameter will tend to produce smaller posterior standard deviations of ability. We feel that this explains why the standard deviations did not increase as expected.

In order to put both sets of estimated item paramters on a common scale and make them more comparable, we propose the following transformation. Define $\frac{1}{k}$

fine
$$\sigma = \left(\begin{bmatrix} k \\ \Pi \\ j = 1 \end{bmatrix} \right)^{1/k},$$

$$\mu = -\sigma \frac{\sum_{j=1}^{k} \beta}{k}, \qquad (19)$$

$$\alpha_{j}^{*} = \frac{\alpha_{j}}{\sigma}$$
, and

$$\beta_{j}^{*} = \mu + \sigma \beta_{j}, \quad j = 1, \ldots, k.$$

The corresponding transformation for the ability parameter is

$$\theta_{i}^{*} = \mu + \sigma\theta_{i}. \tag{20}$$

We refer to this process as rescaling and it is easily seen that $\lim_{j=1}^k \alpha_j^* = 1$ and $\lim_{j=1}^k \beta_j^* = 0$. Under this transformation we now have

$$E(\theta_{i}^{*}|\underline{y}) \approx \mu + \sigma E(\theta_{i}|\underline{y})$$
and
$$SD(\theta_{i}^{*}|\underline{y}) \approx \sigma SD(\theta_{i}|\underline{y}).$$
(21)

One reason for putting the restriction on the item parameter space rather than on the abilities is because the item parameters are common to both sets while the ability parameters in the two sets are not identical, except for the first 100. Upon applying the rescaling to the two sets separately, the rescaled posterior standard deviations for n = 100 were .6 to 14 percent larger than those found using n = 400. This can be seen graphically in Figure 4.

Discussion

The approximation to the posterior mean and variance of the ability parameter proposed here is relatively simple to compute once the posterior mode and approximate covariance matrix of the item parameters are available. Other approximations, which do not appear as readily adaptable to ability estimation, yet deserve further study, include those by Leonard (1982) and Tierney and Kadane (1986). We now briefly examine the computational requirements of these approximations for evaluating the posterior mean 9 of an individual with response x who appears subsequent to the calibration based on n individuals with data Y.

For the Tierney and Kadane approximation define

$$\ell(\xi) = \prod_{i=1}^{m} \int_{j=1}^{k} P(y_{ij} | \theta_i, \xi_j) p(\theta_i) d\theta_i, \qquad (22)$$

$$L(\xi) = m^{-1} \log \pi(\xi) \ \ell(\xi), \tag{23}$$

and

$$L^{*}(\xi) = m^{-1} \log E(\theta | X, \xi) \pi(\xi) \varepsilon(\xi), \qquad (24)$$

where $\pi(\xi)$ is the prior for ξ , m = n+1, $\theta_{n+1} = \theta$ and $(y_{n+1,1}, \dots, y_{n+1,k})^T$ = x so that the posterior expectation of θ can be expressed as

$$E(\theta | \underline{y}, \underline{x}) = \frac{\int e^{mL^{\bullet}(\underline{\xi})} d\underline{\xi}}{\int e^{mL(\underline{\xi})} d\underline{\xi}}.$$
 (25)

Suppose ξ' and ξ^* are the modes of L and L* and ξ' and ξ^* are minus the inverse Hessians of L and L* evaluated at ξ' and ξ^* , respectively. Under regularity conditions, the posterior expectation may be approximated, according to Tierney and Kadane by

$$\widetilde{\theta} = \frac{\left|\sum_{i=1}^{\infty}\right|^{\frac{1}{2}}}{\left|\sum_{i=1}^{\infty}\right|^{\frac{1}{2}}} = \exp\left\{m\left[L^{*}\left(\underbrace{\xi^{*}}\right) - L\left(\underbrace{\xi^{*}}\right)\right]\right\}. \quad (26)$$

One of the requirements is that $E(\theta \mid \mathbf{x}, \boldsymbol{\xi})$ be a positive valued function of $\boldsymbol{\xi}$. But as pointed out by R.D. Bock (personal communication, April 29, 1986) since the origin of θ is arbitrary we may take it to be a large positive number, such as 5, and if necessary truncate the prior of θ to make $E(\theta \mid \mathbf{x}, \boldsymbol{\xi})$ positive.

Although the computation of $(\frac{\hat{\zeta}}{\hat{\zeta}}', \underline{\Sigma}')$ can be performed using the EM algorithm as shown in Tsutakawa and Lin (1986), the computation of $\underline{\xi}^*$ can be a major problem, since, unlike the computation for ξ' , the maximization of $L^*(\underline{\xi})$ requries working with all item parameters simultaneously.

For the Leonard approximation, we begin with the joint posterior of (θ,ξ) ,

$$p(\theta, \xi | \mathbf{x}, \mathbf{y}) = p(\mathbf{x} | \theta, \xi) p(\theta) p(\xi | \mathbf{y})$$
 (27)

For fixed θ , let $\overset{\wedge}{\xi}_{\theta}$ be the mode of ξ and $\overset{\wedge}{\aleph}_{\theta}$ the negative inverse Hessian of the log of (27) with respect to ξ evaulated at $\overset{\wedge}{\xi}_{\theta}$. Then, under regularity conditions, the marginal posterior pdf of θ is approximated by

Ability Estimation

$$p(\theta | \mathbf{x}, \mathbf{y}) = (2\pi)^{q/2} p(\theta, \mathbf{\xi}_{\theta} | \mathbf{x}, \mathbf{y}) / |\mathbf{R}_{\theta}|^{5}, \quad (28)$$

where q is the dimension of ξ . Following Leonard and Novick (1986), the posterior moments of θ may be computed from (28) using numerical integration.

In practice the major difficulty would be in computing $\hat{\xi}_0$. The EM algorithm can again be used, with some modification of the procedure in Tsutakawa and Lin (1986). However since the computation of each ξ_0 will require as much effort as computing and a fair number of $(\hat{\xi}_0,|\hat{\chi}_0|)$ pairs will be needed for each individual, without some modification, Leonard's approximation.

In practice the simplicity of our approximation is obtained by separating the item parameter and ability estimation.

Although there is some loss of information from not using all current data successively update the posterior distribution of the item parameters, (i) the extensive computation is completed at the calibration stage and (ii) ability measurement is done uniformly and without excessive computation for subsequent examinees.

References

- Anderson, T.W. (1958). An Introduction to Multivariate Statistical Analysis (1st ed.). New York: John Wiley and Sons.
- Bock, R.D. & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of the EM algorithm. <u>Psychometrika</u>, <u>46</u>, 443-459.
- Cox, D.R. & Hinkley, D.V. (1974). Theoretical Statistics. London: Chapman & Hall.
- Deeley, J.J. & Lindley, D.V. (1981). Bayes empirical Bayes.

 Journal of the American Statistical Association, 76, 833-841.
- Leonard, T. (1982). Comment on "A simple predictive density function", by M. Lejeune & G.D. Faulkenberry. <u>Journal of the American Statistical Association</u>, 77, 657-658.
- Leonard, T. & Novick, M.R. (1986). Bayesian full rank marginalization for two-way contingency tables. <u>Journal of Educational</u> <u>Statistics</u>, <u>11</u>, 33-56.
- Lindley, D.V. (1980). Approximate Bayesian methods. <u>Trabajos</u> Estadistica, 31, 223-237.
- Lord, F.M. (1983a). Unbiased estimators of ability parameters, of their variance and of their parallel-forms reliability. Psychometrika, 48, 233-245.
- Lord, F.M. (1983b). Statistical bias in maximum likelihood eximators of item parameters. <u>Psychometrika</u>, <u>48</u>, 425-435.
- Mosteller, F. & Wallace, D.L. (1964). <u>Inference and Disputed</u>
 <u>Authorship: The Federalist</u>. Reading, MA: Addison Wesley.
- Rigdon, S.E. & Tsutakawa, R.K. (1983). Estimation in latent trait models. Psychometrika, 48, 567-574.
- Tierney, L. & Kadane, J.B. (1986). Accurate approximations for posterior moments and marginal densities. <u>Journal of the American Statistial Association</u>, <u>81</u>, 82-86.
- Tsutakawa, R.K. (1984). Estimation of two-parameter logistic item response curves. Journal of Educational Statistics, 9, 263-267.
- Tsutakawa, R.K. & Lin, H.Y. (1986). Bayesian estimation of item response curves. Psychometrika, 51, 251-267.
- Wingersky, M.S., Barton, M.A., & Lord, F.M. (1982). <u>Logist</u>
 User's Guide. Princeton, NJ:Educational Testing Service.

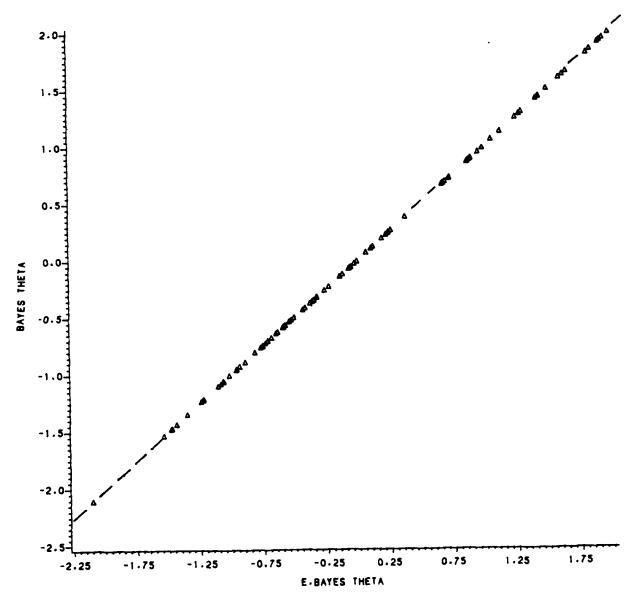


Figure 1. Bayes vs. Empirical Bayes θ .

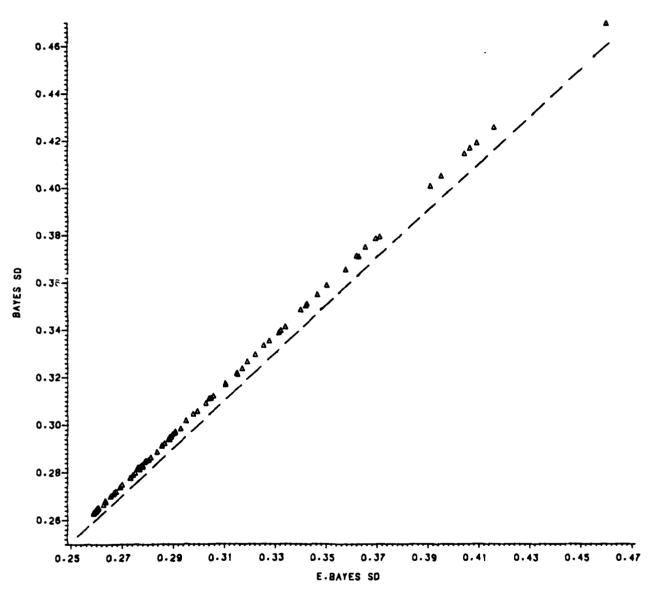


Figure 2. Bayes vs. Empirical Bayes Standard Deviation.

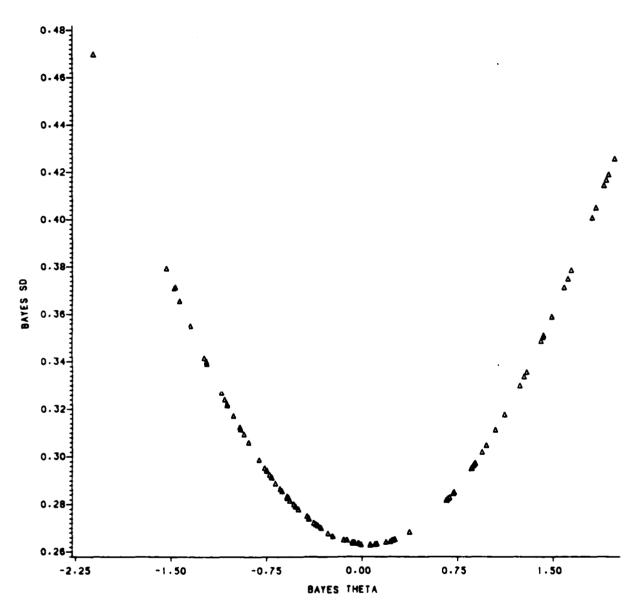


Figure 3. Bayes Standard Deviation vs. Bayes %.

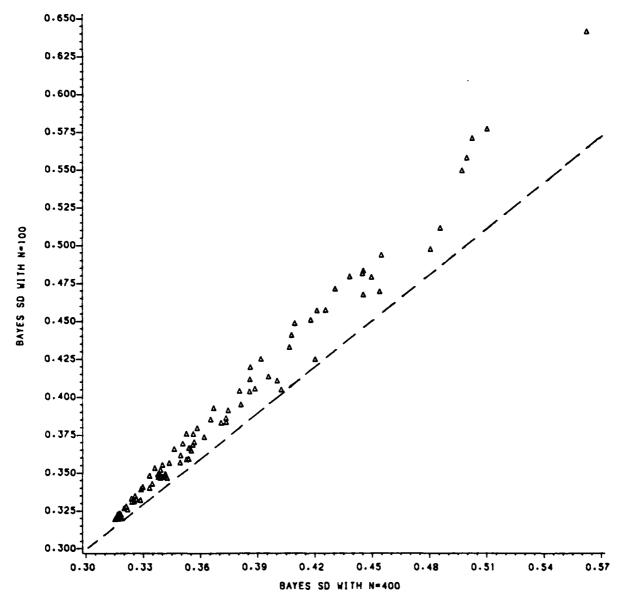


Figure 4. Rescaled Bayes Standard Deviations for N = 100 vs. 400

Dr. Terry Ackerman American College Testing Programs P.O. Box 168 Iowa City, IA 52243

Dr. Robert Ahlers Code N711 Human Factors Laboratory Naval Training Systems Center Orlando, FL 32613

Dr. James Algina University of Florida Gainesville, FL 32605

Dr. Erling B. Andersen Department of Statistics Studiestraede 6 1455 Copenhagen DENMARK

Dr. Eva L. Baker
UCLA Center for the Study
of Evaluation
145 Moore Hall
University of California
Los Angeles, CA 90024

Dr. Isaac Bejar Educational Testing Service Princeton, NJ 08450

Dr. Menucha Birenbaum School of Education Tel Aviv University Tel Aviv, Ramat Aviv 69978 ISRAEL

Dr. Arthur S. Blaiwes Code N711 Naval Training Systems Center Orlando, FL 32813

Dr. Bruce Bloxom
Defense Manpower Data Center
550 Camino El Estero,
Suite 200
Monterey, CA 93943-3231

Dr. R. Darrell Bock University of Chicago NORC 6030 South Ellis Chicago, IL 60637

Cdt. Arnold Bohrer
Sectie Psychologisch Onderzoek
Rekruterings-En Selectiecentrum
Kwartier Koningen Astrid
Bruijnstraat
1120 Brussels, BELGIUM

Dr. Robert Breaux Code N-095R Naval Training Systems Center Orlando, FL 32813

Dr. Robert Brennan American College Testing Programs P. O. Box 168 Iowa City, IA 52243

Dr. Lyle D. Broemeling ONR Code 1111SP 800 North Quincy Street Arlington, VA 22217

Mr. James W. Carev Commandant (G-PTE) U.S. Coast Guard 2100 Second Street, S.W. Washington, DC 20593

Dr. James Carlson American College Testing Program P.O. Box 168 Iowa City, IA 52243

Dr. John B. Carrell 409 Elliott Rd. Chapel Hill, NC 27514

Dr. Robert Carroll OP 01B7 Washington, DC 20370

Mr. Raymond E. Christal AFHRL/MOE Brooks AFB, TX 78235

Dr. Norman Cliff
Department of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

Director,
Manpower Support and
Readiness Program
Center for Naval Analysis
2000 North Beauregard Street
Alexandria, VA 22311

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Guincy Street
Arlington, VA 22217-5000

Dr. Hans Crombag University of Leyden Education Research Center Boerhaavelaan 2 2334 EN Leyden The NETHERLANDS

Mr. Timothy Davey University of Illinois Educational Psychology Orbana, IL 61801

Or. Doug Davis
Crief of Naval Education
and Training
Naval Air Station
Fensacola, FL 32508

Or. Dattprasad Divgi Center for Naval Analysis 4401 Ford Avenue F.O. Box 16268 Alexandria, VA 22302-0268

Dr. Hei-Fi Dong Ball Foundation 800 Roosevelt Road Building C, Suite 206 Glen Ellyn, IL 60137 Dr. Fritz Drasgow University of Illinois Department of Psychology 603 E. Daniel St. Champaign, IL 61820

Defense Technical Information Center Cameron Station, Bidg 5 Alexandria, VA 22314 Attn: TC (12 Copies)

Dr. Stephen Dunbar Lindquist Center for Measurement University of Iowa Iowa City, IA 52242

Dr. James A. Eartes Air Force Human Resources Lab Brooks AFB, TX 78235

Dr. Kent Eaton Army Research Institute 5001 Eisenhower Avenue Alexandria, VA 22333

Dr. John M. Eddins
University of Illinois
252 Engineering Research
Laboratory
103 South Mathews Street
Urbana, IL 61601

Dr. Susan Embretson University of Kansas Psychology Department 426 Fraser Lawrence, KS 66045

Dr. Benjamin A. Fairbank Performance Metrics, Inc. 5825 Callaghan Suite 225 San Antonio, TX 78228

Dr. Pat Federico Code 511 NPRDC San Diego, CA 92152-6800

Or. Leonard Feldt Lindquist Center for Measurement on versity of Iowa Iowa City, IA 52242

Dr. Richard L. Ferguson American College Testing Program P.J. Box 168 Iswa City, IA 52240

Ir. Gerhard Fischer
L ebiggasse 5:3
A 1010 Vienna
AUSTRIA

Ir. Myron Fischi Army Research Institute 5001 Eisenhower Avenue Alexandria, VA 22333

Prof. Donald Fitzgerald of versity of New England Department of Psychology Armidake, New South Wales 2351 4,3764,14

Mil Paul Pole. Naul Personnel R&D Center nam Diego, CA 32152-6800

or. Aftred R. Fregly do Palty do Ong APB, DC 20332

ur. Futert D. Gittons
on versity of Illinois-Chicago
on Elv 6998
on ay , 16 63660

cr. January 6 fromd in versity of Massachusetts autopinof Education Amberst, MA 01003

Or. Robert Glaser
Learning Research
8 Development Center
on versity of Pittsburgh
3939 O Hara Street
Fittsburgh, PA 15260

Dr. Bert Green
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dipl. Pad. Michael W. Habon Universitat Dusseldorf Erziehungswissenschaftliches Universitatsstr. 1 D-4000 Dusseldorf 1 WEST GERMANY

Dr. Ronald K. Hambleton
Prof. of Education & Psychology
University of Massachusetts
 at Amherst
Hills House
Amherst, MA 01003

Dr. Delwyn Harnisch University of Illinois 51 Gerty Drive Champaign, IL 61820

Ms. Rebecca Hetter Navy Personnel R&D Center Code 62 San Dregu, CA 92152-6800

Er. Paul W. Holland Educational Testing Service Rosedale Road Frinceton, NJ 08541

Prof. Lutz F. Hornke Institut für Psychologie RWTH Aachen Jaegerstrasse 17/19 D-5100 Aachen WEST GERMANY

Dr. Paul Horst 677 G Street, #184 Chula Vista, CA 90010

Mr. Dick Hoshaw
OP-135
Arlington Annex
Ruom 2834
Washington, DC 20350

Dr. Lloyd Humphreys University of Illinois Department of Psychology 603 East Daniel Street Champaign, IL 61820

Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

Dr. Huynh Huynh College of Education Univ. of South Carolina Columbia, SC 29208

Dr. Robert Jannarone Department of Psychology University of South Carolina Columbia, SC 29208

Dr. Dennis E. Jennings Department of Statistics University of Illinois 1409 West Green Street Urbana, IL 61801

Dr. Douglas H. Jones Thatcher Jones Associates P.O. Box 6640 10 Trafalgar Court Lawrenceville, NJ 08648

Dr. Milton S. Katz Army Research Institute 5001 Eisenhower Avenue Alexandria, VA 22333

Prof. John A. Keats
Department of Psychology
University of Newcastle
N.S.W. 2308
AUSTRALIA

Dr. G. Gage Kingsbury
Portland Public Schools
Research and Evaluation Department
501 North Dixon Street
P. O. Box 3107
Portland, OR 97209-3107

Dr. William Koch University of Texas-Austin Measurement and Evaluation Center Austin, TX 78703

Dr. Leonard Kroeker Navy Personnel R&D Center San Diego, CA 92152-6800

Dr. Daryll Lang Navy Personnel R&D Center San Diego, CA 92152-6800

Dr. Thomas Leonard University of Wisconsin Department of Statistics 1210 West Dayton Street Madison, WI 53705

Dr. Michael Levine Educational Psychology 210 Education Bldg. University of Illinois Champaign, IL 61801

Dr. Charles Lewis Educational Testing Service Princeton, NJ 08541

Dr. Robert Linn College of Education University of Illinois Urbana, IL 61801

Dr. Robert Lockman Center for Naval Analysis 4401 Ford Avenue P.O. Box 16268 Alexandria, VA 22302-0268

Dr. Frederic M. Lord Educational Testing Service Princeton, NJ 08541

Dr. James Lumsden
Department of Psychology
University of Western Australia
Nedlands W.A. 6009
AUSTRALIA

Dr. Milton Maier Center for Naval Analysis 4401 Ford Avenue P.O. Box 16268 Alexandria, VA 22302-0268

Dr. William L. Maloy Chief of Naval Education and Training Naval Air Station Pensacola, FL 32508

Dr. Gary Marco Stop 31-E Educational Testing Service Princeton, NJ 08451

Dr. Clessen Martin Army Research Institute 5001 Eisenhower Blvd. Alexandria, VA 22333

Dr. James McBride
Psychological Corporation
c/o Harcourt, Brace,
 Javanovich Inc.
1250 West 6th Street
San Diego, CA 92101

Dr. Clarence McCormick HQ, MEPCOM MEPCT-P 2500 Green Bay Road North Chicago, IL 60064

Dr. Robert McKinley Educational Testing Service Princeton, NJ 08541

Dr. James McMichael Technical Director Navy Personnel R&D Center San Diego, CA 92152

Dr. Barbara Means Human Resources Research Organization 1100 South Washington Alexandria, VA 22314

Dr. Robert Mislevy Educational Testing Service Princeton, NJ 08541 Dr. William Montague NPRDC Code 13 San Diego, CA 92152-6800

Ms. Kathleen Moreno Navy Personnel R&D Center Code 62 San Diego, CA 92152-6800

Headquarters, Marine Corps Code MPI-20 Washington, DC 20380

Dr. W. Alan Nicewander University of Oklahoma Department of Psychology Oklahoma City, OK 73069

Deputy Technical Director NPRDC Code 01A San Diego, CA 92152-6800

Director, Training Laboratory, NPRDC (Code 05) San Diego, CA 92152-6800

Director, Manpower and Personnel Laboratory, NPRDC (Code 06) San Diego, CA 92152-6800

Director, Human Factors & Organizational Systems Lab, NPRDC (Code 07) San Diego, CA 92152-6600

Fleet Support Office, NPRDC (Code 301) San Diego, CA 92152-6800

Library, NPRDC Code P201L San Diego, CA 92152-6800

Commanding Officer, Naval Research Laboratory Code 2627 Washington, DC 20390

Dr. Harold F. O'Neil, Jr.
School of Education - WPH 801
Department of Educational
Psychology & Technology
University of Southern California
Los Angeles, CA 90089-0031

Dr. James Olson WICAT, Inc. 1875 South State Street Orem, UT 84057

Office of Naval Research, Code 1142PT 800 N. Quincy Street Arlington, VA 22217-5000 (6 Copies)

Office of Naval Research.
Code 125
800 N. Quincy Street
Arlington, VA 22217-5000

Assistant for MPT Research, Development and Studies OP 0167 Washington, DC 20370

Dr. Judith Orasanu Army Research Institute 5001 Eisenhower Avenue Alexandria, VA 22333

Dr. Jesse Orlansky Institute for Defense Analyses 1801 N. Beauregard St. Alexandria. VA 22311

Dr. Randolph Park Army Research Institute 5001 Eisenhower Blvd. Alexandria, VA 22333

Wayne M. Patience American Council on Education GED Testing Service, Suite 20 One Dupont Circle, NW Washington, DC 20036 Dr. James Paulson Department of Psychology Portland State University P.O. Box 751 Portland, OR 97207

Administrative Sciences Department, Naval Postgraduate School Monterey, CA 93940

Department of Operations Research, Naval Postgraduate School Monterey, CA 93940

Dr. Mark D. Reckase ACT P. O. Box 168 Iowa City, IA 52243

Dr. Malcolm Ree AFHRL/MP Brooks AFB, TX 78235

Dr. Barry Riegelhaupt HumRRO 1100 South Washington Street Alexandria, VA 22314

Dr. Carl Ross CNET-PDCD Building 90 Great Lakes NTC, IL 60086

Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

Dr. Fumiko Samejima Department of Psychology University of Tennessee Knoxville, TN 37916

Mr. Drew Sands NPRDC Code 62 San Diego, CA 92152-6600

Lowell Schoer
Psychological & Quantitative
Foundations
College of Education
University of Iowa
Iowa City, IA 52242

Dr. Mary Schratz Navy Personnel R&D Center San Diego. CA 92152-6800

Dr. Dan Segall Navy Personnel R&D Center San Diego, CA 92152

Dr. W. Steve Sellman DASD(MRA&L) 2B269 The Pentagon Washington, DC 20301

Dr. Kazuo Shigemasu 7-9-24 Kugenuma-Kaigan Fujusawa 251 JAPAN

Dr. William Sims Center for Naval Analysis 4401 Ford Avenue P.O. Box 16268 Alexandria, VA 22302-0268

Dr. H. Wallace Sinaiko Manpower Research and Advisory Services Smithsonian Institution SGI North Pitt Street Alexandria, VA 22314

Dr. Richard E. Snow Department of Psychology Stanford University Stanford, CA 94306

Dr. Richard Sorensen Navy Personnel R&D Center San Diego, CA 92152-6800

Dr. Paul Speckman University of Missouri Department of Statistics Columbia, MO 65201

Dr. Judy Spray ACT F.O. Box 168 Iowa City, IA 52243

Dr. Martha Stocking Educational Testing Service Princeton, NJ 08541 Dr. Peter Stoloff Center for Naval Analysis 200 North Beauregard Street Alexandria, VA 22311

Dr. William Stout University of Illinois Department of Mathematics Urbana, IL 61801

Maj. Bill Strickland AF/MPXOA 4E166 Pentagon Washington, DC 20330

Dr. Hariharan Swaminathan Laboratory of Psychometric and Evaluation Research School of Education University of Massachusetts Amherst, MA 01003

Mr. Brad Sympson Navy Personnel R&D Center San Diego, CA 92152-6800

Dr. John Tangney AFOSR/NL Bolling AFB, DC 20332

Dr. Kikumi Tatsuoka CERL 252 Engineering Research Laboratory Urbana, IL 61601

Dr. Maurice Tatsucka 220 Education Bldg 1310 S. Sixth St. Champaign, IL 61820

Dr. David Thissen Department of Psychology University of Kansas Lawrence, KS 66044

Mr. Gary Thomasson University of Illinois Educationa! Psychology Champaign, IL 61820

Dr. Robert Tsutakawa University of Missouri Department of Statistics 222 Math. Sciences Bldg. Columbia, MO 65211

Dr. Ledyard Tucker University of Illinois Department of Psychology 603 E. Daniel Street Champaign, IL 61820

Dr. Vern W. Urry Personnel R&D Center Office of Personnel Management 1900 E. Street, NW Washington, DC 20415

Dr. David Vale Assessment Systems Corp. 2233 University Avenue Suite 310 St. Paul, MN 55114

Dr. Frank Vicino Navy Personnel R&D Center San Diego, CA 92152-6800

In . moward warner
I wision of Psychological Studies
Educational Testing Service
Princeton, No 63541

in MingaMc aling indquist Center for Measurement University of Iowa Iowa City, IA 52242

Or. Thomas A. Warm Coast Guard Institute P. O. Substation 18 Oklahoma City, OK 73169

Dr. Brian Waters
Program Manager
Manpower Analysis Program
HumRRO
1100 S. Washington St.
Alexandria, VA 22314

Dr. David J. Weiss NG60 Elliott Hall University of Minnesota 75 E. River Road Minneapolis, MN 55455

Dr. Ronald A. Weitzman NPS, Code 54Wz Monterey, CA 92152-6800

Major John Welsh AFHRL/MOAN Brooks AFB, TX 76223

Dr. Douglas Wetzel Code 12 Navy Personnel R&D Center San Diego, CA 92152~6600

Dr. Rand R. Wilcox University of Southern California Department of Psychology Los Angeles, CA 90007

German Military Representative ATTN: Wolfgang wildegrube Streit-Paefiteant D-5300 Bonn 2 4000 Brandywine Street, NW Washington, DC 20016

Department of Educational Psychology University of Illinois Urbana, IL 61801

Dr. Hilda Wing
Psychological Corporation
c/o Harcourt, Brace,
Javanovich Inc.
1250 West 6th Street
San Diego, CA 92101

Dr. Martin F. Wiskoff Navy Personnel R & D Center San Diego, CA 92152-6800

Mr. John H. Wolfe Navy Personnel R&D Center San Diego, CA 92152-6600

Dr. George Wong Biostatistics Laboratory Memorial Sloan-Kettering Cancer Center 1275 York Avenue New York, NY 10021

Dr. Wallace Wulfeck, III Navy Personnel R&D Center San Diego, CA 92152-6800

Dr. Wendy Yen CTB/McGraw Hill Del Monte Research Park Monterey, CA 93940

Dr. Joseph L. Young
Memory & Cognitive
Processes
National Science Foundation
Washington, DC 20550

